Cognitive Development 60 (2021) 101125

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Perceiving size through sound in sighted and visually impaired children

Luigi F. Cuturi[a,*], Giulia Cappagli[a], Alessia Tonelli[a], Elena Cocchi[b], Monica Gori[a]

[a] *Unit for Visually Impaired People, Istituto Italiano di Tecnologia, Genoa, Italy*
[b] *Istituto David Chiossone, Genoa, Italy*

ARTICLE INFO

Keywords:
Blindness
Childhood blindness
Size perception
Crossmodal correspondences
Multisensory mechanisms

ABSTRACT

Associations between sensory features of different natures are defined as crossmodal correspondences. In the context of size perception, low pitch sound frequencies are often associated with larger objects and high pitch with smaller objects. Here we investigate such crossmodal correspondences in sighted and visually-impaired children. In Experiment 1, after listening to sounds (250–5000 Hz pure tones), children aged 6–11 years were asked to draw a circle "as big as the sound was". In Experiment 2, children aged 6–14 years who were blind or had low vision performed a similar task. In accordance with previous research, we observed that the circle size drawn depends on participants' age and we confirm the presence of pitch-size associations in sighted children. In visually-impaired children, such associations are influenced by residual vision, suggesting an anchoring of size perception to level of residual vision. These results reveal novel dynamics underlying the advancing of visual loss and the emergence of compensatory mechanisms in childhood.

1. Introduction

During childhood, drawing is one of the first and most important acquired ways to represent the surrounding world in all aspects, including objects of different sizes. Drawing ability changes during development; specifically, in the context of drawn size. For instance, in representing details of an object, children tend to exaggerate drawing the size of smaller features (Silk & Thomas, 1988). Drawing ability depends on two main factors: the spatial properties of the object to be drawn, namely an aggregation factor where figures need to fit into the available drawing space, and the complexity of the spatial axes system, such as the transfer of 3D-objects in a bidimensional drawing (Lange-Küttner, 1997, 2004, 2009). These two factors have been identified as the most influential across development (Lange-Küttner, 1997, 2004). Specifically, overall size reduction in a complex drawing such as that of a human figure has been observed. In this context, object-related properties mostly influenced children aged 7–9 years, whereas older children were mostly influenced by spatial-axes systems (Lange-Küttner, 1997, 2004, 2009).

To draw an object, physical properties such as its size or shape must be perceptually encoded by the child through the available sensory channels. Size perception is fundamental for developing abilities such as object discrimination (Gori, Giuliana, Sandini, & Burr, 2012) and solving mathematical problems related to magnitude comparison (Fooks, Hadad, & Rubinsten, 2021). Substantial evidence indicates that touch is the most robust sense used to perceive the size of an object. For instance, it has been shown that touch dominates both vision (Gori, Del Viva, Sandini, & Burr, 2008) and hearing (Petrini, Remark, Smith, & Nardini, 2014) in determining

* Correspondence to: Unit for Visually Impaired People, Istituto Italiano di Tecnologia, Via Enrico Melen, 83, Genoa, Italy.
E-mail address: Luigi.Cuturi@iit.it (L.F. Cuturi).

https://doi.org/10.1016/j.cogdev.2021.101125
Received 21 December 2020; Received in revised form 20 September 2021; Accepted 24 September 2021
Available online 1 October 2021

the dimensions of an object in early development, indicating that children heavily rely on touch for size estimation. Such findings suggest that there might exist a process of sensory calibration for size estimation, by which touch calibrates vision and hearing during development (Gori, 2015; Gori et al., 2008; Petrini et al., 2014). However, size can also be conveyed through audition by exploiting multisensory associations known as crossmodal correspondences (Parise, 2015; Spence, 2011). For instance, visual size estimation can be biased by auditory pitch; namely, low sound frequencies can increase perceived size, and high tone frequencies can decrease perceived size (Evans & Treisman, 2011; Parise & Spence, 2009; Tonelli, Cuturi, & Gori, 2017). The emergence of such crossmodal correspondences is unrelated to cultural connotations and likely arises from prior experience and interaction with the surrounding environment (Eitan & Timmers, 2010). In this context, vision certainly plays a strong role in conveying size information about objects we interact with. For instance, the association of pitch with visually-presented bouncing balls of different sizes is apparent at 30 months of age (Mondloch & Maurer, 2004). Similarly, studies on infants showed that the association between size and pitch is evident at 4 months of age (Dolscheid, Hunnius, Casasanto, & Majid, 2014; Peña, Mehler, & Nespor, 2011) and 6 months with speech-independent auditory stimuli (Fernández-Prieto, Navarra, & Pons, 2015). Nevertheless, multisensory experience and top-down processes seem to be crucial factors in guiding the emergence of such perceptual phenomena across development (Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012; Orchard-Mills, Van der Burg, & Alais, 2013). While an early onset of cross-modal mapping has been clearly observed in infancy, cross-modal mapping in middle childhood most frequently is investigated in reading research (e.g. print-sound mapping, see Nag & Snowling, 2013) and in research on modality-specific working memory (e.g. auditory loop and visuospatial sketchpad, see Vuontela, Steenari, & Koivisto, 2003).

Crossmodal correspondences such as those relative to size can be interpreted in the context of supramodal processes that convey the object's properties, independently of encoding sensory modalities. The ATOM theory, proposed by Walsh (2003) suggests that magnitude estimation derives from common processing mechanisms that link different dimensions (e.g. space and time or number and time), which might involve different sensory modalities in the processing of size estimation. In general, the origin of crossmodal correspondences seems to be related to redundancies and associations present in the surrounding environment. For instance, objects small in size tend to produce sounds with a higher pitch compared to large objects that instead produce lower pitch sounds. This is true in the context of musical instruments for which the physical principle that guides the instrument's pitch extension naturally follows the correspondence between size and pitch (e.g., the violin family, ranging from the smaller violin to the larger double bass, or the saxophone family from the smaller alto saxophone to the larger bass saxophone). Generally, not only objects but also living beings such as humans, and in many cases animals show examples of crossmodal correspondences relative to pitch, with for instance children producing higher pitch voices compared to adults who produce lower pitch voices. The exposure to such multisensory redundancy may underlie the origin of crossmodal correspondences which seems to appear quite early during development (Ozturk, Krehm, & Vouloumanos, 2013; Peña et al., 2011). Moreover, in the context of multisensory integration, the combination of redundant sensory information increases precision in perception (Ernst & Banks, 2002); in this sense, additional size encoding via auditory cues may improve the ability to discriminate objects of different sizes when both visual and auditory cues are present. At the same time, auditory cues such as pitch can improve perceived relative timing of visual and auditory events in a size estimation task (Jaekl, Soto-Faraco, & Harris, 2012), indicating that auditory processing can be informative about environmental properties even if unconventionally related to audition such as size estimation.

While it is known that size perception is typically coded within the haptic modality and that touch calibrates vision for size perception (Gori et al., 2008), less is known about how individuals deprived of vision encode size information through audition. Since audition is not intrinsically related to size, it can be argued that haptic experience might be enough to gain such crossmodal correspondences. Indeed, it has been shown that pitch-size correspondence does not require visual experience (Hamilton-Fletcher *et al.*, 2018; Scheller, Proulx, De Haan, Dahlmann-Noor, & Petrini, 2021); however, how such association emerges across the lifespan of a visually-impaired child remains unclear. The ability to use non-visual senses to process object features is fundamental to build a compensatory strategy for lack of vision. Studies in children have revealed that a lack of vision severely compromises the ability to perform tasks typically conveyed within the visual modality, such as orientation discrimination and spatial localization. For instance, visually-impaired children show impairments in the ability to judge objects' orientation with the haptic modality (Gori, Sandini, Martinoli, & Burr, 2010), stimuli distance (Cappagli, Cocchi, & Gori, 2015) and localization (Cappagli & Gori, 2016) within the auditory modality. In this context, it has been shown that in the absence of visual experience, the interaction of auditory and tactile information is reduced (Hötting & Röder, 2004), likely because vision acts as a binding sensory modality to provide multisensory integration. In this sense, complete or partial lack of visual information might differentially affect the development of crossmodal correspondence (e.g. audio-tactile) during childhood and especially in relationship to severity of visual impairment. Surprisingly, associations related to crossmodal correspondences, such as the ability to perceive size through sounds, have not been well investigated in visually-impaired children. Such a perceptual phenomenon might represent the expression of a natural compensatory mechanism for processing size through the remaining senses in case of visual loss.

In a recent study (Cuturi, Tonelli, Cappagli, & Gori, 2019), the authors tested crossmodal correspondences across childhood by asking children to identify the shape corresponding to different sound frequencies. The results indicate a developmental trend specific for the use of intermediate sizes in response to intermediate sound frequencies. According to these findings, older children are better able to take advantage of such precise and fine-grained stimuli compared to younger peers. However, an identification task may contain methodological limitations on the investigation of crossmodal correspondences across childhood. Above all, children are presented with a scale of stimuli, thus their choice, though being free, relies on the presentation of options to be chosen (i.e. bottom-up process) rather than their own production of response to the auditory stimuli (i.e. top-down process). In contrast, a drawing task would represent a methodological choice ideal for testing the spontaneous top-down occurrence of crossmodal correspondences across childhood. Moreover, scientific evidence indicates that motor responses in an egocentric frame of reference are more accurate than conscious responses such as verbal reports (Creem & Proffitt, 2001; Wraga, Creem, & Proffitt, 2000). To assess Audio-visual cross-modal correspondences related to size perception in adults, in a recent study (Tonelli et al., 2017) using a "listen-and-draw" task, participants listened to several sounds differing in frequency and drew a circle as big as they perceived sound using a computer mouse. This study showed that pure tones ranging from low to high frequency were associated with drawn size. In the present study, we use an adapted version of this task to test: a) whether development-related differences in the association between sound pure tone frequency and visual size emerge in sighted children between ages 6 and 11 (Study 1) and b) how visual impairment affects the development of such association across childhood (Study 2). Concerning sighted children, we expect that the association of pure tones with size (crossmodal correspondences) emerge around age 6 and that development fosters finer association of intermediate sound frequency and size as previously observed (Cuturi et al., 2019). The main interest in this study is to extend previous findings with a rather different task, where children are not constrained to choose from a selection of visual stimuli, but instead are free to draw thus leading to the involvement of top-down rather than bottom-up processes. Moreover, considering developmental changes in drawing size previously reported in the literature (Lange-Küttner, 1997, 2004, 2009), we expect older children to make smaller drawings compared to younger ones. In the case of visually-impaired children, we hypothesize that the ability to perceive size through audition would be reduced in low vision children who anchor their perceptual abilities to the residual vision, whereas more robust sensory-loss compensatory processes would take place in totally blind children.

2. Study 1

2.1. Methods

2.1.1. Participants

In Study 1, we studied how sighted children associate visual size and auditory pure tone frequencies across development. 69 sighted children from age 6–11 participated (72–139 months; 37 males; all right-handed, except 4). The number of participants is chosen based on previous studies that investigated size perception in children (e.g.: Cuturi et al., 2019; Gori et al., 2008, 2012; Petrini et al., 2014). All participants reported having normal hearing.

2.1.2. Stimuli

Auditory stimuli were pure tones of different frequency (250, 500, 1000, 2000, 5000 Hz) generated at a sample rate of 44100 Hz using Audacity software (http://audacity.sourceforge.net/). Stimuli were presented through one loudspeaker (Sony™ - SRS-X11) positioned in front of the participant at a distance of ~25 cm. For all auditory stimuli, sound level plus background noise was ~70 dB. Each auditory stimulus lasted 2 s. Given the high rate of potential task withdrawal in children, in order to guarantee performance comparisons among participants, we pseudo-randomized the presentation order and used this fixed order for all participants. We did not observe task withdrawal. Responses were taken with a touch screen 21.5" monitor (Dell™ SX2210TB; screen resolution: 1920 × 1080; refresh rate: 60 Hz) positioned in front of the participant at ~40 cm. In Study 1, each sound frequency was repeated 10 times, leading to 50 trials in total.

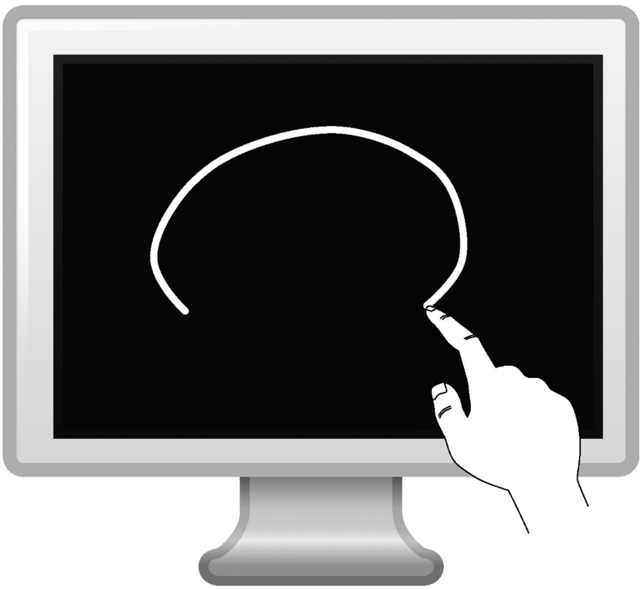

Fig. 1. Reproduction task. Pictorial representation of the task phase in which participants are asked to draw the circle on the touch monitor.

2.1.3. Task

Data were collected in person. The task consisted of drawing a circle in response to the sound stimulus. In detail, after listening to the sound presented, participants were asked to draw a circle "as big as the sound" on a touch screen by using the index finger of their dominant hand. In Study 1, the drawn shape was visible during the drawing phase as a white line on a black background (Fig. 1); when finished with the drawing, the drawn shape disappeared, and participants were asked to rate from 1 to 10 the size of drawn shape (from 1 = very small to 10 = very large). The evaluation of size can be made by either judging object dimension in non-numerical terms, such as on a continuum from “small” to “big”, or by indicating how big an object is using a given numerical scale. In the context of numbers representation, developmental studies have shown that such representation is unevenly spaced in children, with smaller numbers more spaced out and larger numbers more compressed (e.g., Booth & Siegler, 2006). However, studies on crossmodal correspondences have mostly investigated this phenomenon with no involvement of numerical estimates by adopting extreme stimuli on the size continuum that could not disentangle the relative versus absolute nature of such perceptual phenomena (Spence, 2019). In this context, the interrelationship between the development-related differences in numbers representation and the emergence of crossmodal correspondences may represent an influential factor in the context of size estimation and representation. To ensure such relationships were present in the children who participated in the study, we additionally asked participants to rate the size of the drawn figure on a scale from 1 to 10. We included this additional task as a control procedure to check whether children were consistent in the association between drawn size and a symbolic representation such as numbers. In order to make the rating understandable to all children, the experimenter showed them two drawn circles resembling extreme size levels associated with a rating equal to 1 (smallest) and the size associated with the rating equal to 10 (biggest). The concept of a continuum between the two extreme ratings was clearly explained prior to the test: Participants were first asked to count from 1 to 10 and then informed that they could use units ranging from 1 to 10.

2.2. Data analyses

We analyzed participants’ performance, considering the drawn size estimate in cm^2 in response to the sound stimulus. First, we calculated the drawn area using the MATLAB® function “polyarea.m”, which calculates the area of a 2-D polygon. In the case of open drawings, when the participant did not close the figure, the function takes into account the minimum distance between the endpoints of the drawn line. For each repetition, we used the size of the drawing in cm^2 as the measurement of the estimated size that corresponds to the auditory stimulus, hereafter “size estimate”. Considering there are no boundaries (except for the monitor frame) to limit the drawing corresponding to the presented sound frequency, high variability between participants is expected. Therefore, we used a z-transformation on size estimate in cm^2 on a participant-by-participant basis, that is we calculated the z-transformed size based on all responses of each individual participant. This choice grounds on the intention to analyze performance by reducing the impact of interindividual variability. For each calculated size estimate, we used an outlier exclusion criterion applied in several studies on multisensory perception both in adults (e.g., Tivadar et al., 2019; Tonelli et al., 2017) and across childhood (e.g., Amadeo, Campus, & Gori, 2019; Honoré & Noë, 2016; Wiedenbauer & Jansen-Osmann, 2008), by removing those size estimate values that deviate more than 2 SD from the mean across all measurements for each sound frequency.

We considered the standard deviation across repetitions as a measurement of response variability within-participant, hereafter “variability”.

We used the lme4 package in R to separately fit linear mixed-effects models to size estimate and variability. To evaluate the significance of the model effects, we performed an ANOVA on the model estimates using Kenward-Roger’s degrees of freedom approximation (Luke, 2017). For statistically significant linear models, we calculated estimated marginal means (EMMEANs for the categorical predictor variable, e.g. pure tone frequency) and estimated marginal means of linear trends (EMTRENDs for continuous predictor variable, e.g. age decimal). We report EMMEANs and EMTRENDs with 95% confidence intervals (CI) (Lenth et al., 2021). The EMMEAN demonstrates the mean response of the response variable for the predictor variable, i.e. Pure tone frequency. All statistical analyses were performed using R. Significance was set with an alpha-value of 0.05.

We tested the influence of the between-factor “Age in years” (treated as a continuous variable in decimal values) and the within-factor “Pure tone frequency” (with five levels: 250, 500, 1000, 2000, 5000 Hz) on size estimate. The variability of the participant factor and the trials’ order of presentation were taken into account by modeling them as fixed random effects. The model is thus described by the following formula:

RESPONSE ~ Pure tone frequency x Age in years + (1| Participant) + (1| Trial Order)

where RESPONSE indicates the size estimate for each given trial after outlier exclusion. The EMTREND shows the mean change in the response variable for a unit change in the continuous predictor variable, i.e. Age in years, adjusted for other predictor variables in the model, i.e. each level of the Pure tone frequency factor (i.e. 250, 500, 1000, 2000, 5000 Hz).

We conducted a similar analysis for variability using a linear-mixed effects model described as follows:

VARIABILITY ~ Pure tone frequency x Age in years + (1| Participant)

where VARIABILITY indicates the standard deviation across repetitions for each sound frequency and participant. To evaluate the significance of the model effects, we performed ANOVA on the mixed model using Kenward-Roger’s degrees of freedom approximation (Luke, 2017). Post hoc analysis was conducted by comparing contrasts using the emmeans package (Lenth et al., 2021). With this analysis, we computed and contrasted predicted probability distributions relative to each factor. We considered pairwise comparisons with $P < 0.05$ to be significant (Tukey’s corrected *ps* are reported).

Additionally, we tested the relationship between participants' variability and their size estimate to test whether more variable participants tended to draw bigger or smaller circles in response to each frequency. Specifically, we ran a correlation analysis between participants' variability and averaged size estimate in cm^2 across repetitions for each sound frequency. As data are not normally distributed, we used Spearman's non-parametric correlation analysis.

To test whether participants are able to translate their drawn objects to a numerical scale, for each Pure tone frequency we ran a correlation analysis of the relationship between the size estimate and the number assigned by participants to the response drawing. Each data point corresponds to the response given at an individual trial. We used Spearman's non-parametric correlation analysis.

2.3. Results

The outlier exclusion led us to remove 4.84% of data points for the raw data in cm^2 and 4.84% of data points for the z-transformed data. Overall, the results show that children of all examined ages associate large drawings with low pure tone frequencies and small drawings with high pure tone frequencies (Fig. 2), indicating that such association is present by age 6.

For all data sets, ANOVAs on the linear mixed model for each measurement of size estimation show a significant effect of the Pure tone frequency (raw data in cm^2: F (4, 2594.30) = 66.64; $p < 0.001$; z-transformed data: F (4, 1422.40) = 43.20; $p < 0.001$) and no effect of Age in years (raw data in cm^2: F (1, 67.20) = 0.29; $p = 0.589$; z-transformed data: F (4, 67.70) = 43.20; $p < 0.001$). Post hoc analysis on the raw data in cm^2 and the z-transformed data reveal a similar pattern showing significant differences for all comparisons between Pure tone frequencies (except for the non-significant comparison between 2000 and 5000 Hz; Supplementary Table S1). In the analysis of raw data in cm^2 we observe a significant interaction between Pure tone frequency and Age in years (F (4, 3163.60) = 15.00; $p < 0.001$), not present for the z-transformed data (F (4, 3171.10) = 1.66; $p = 0.154$). Post-hoc analysis shows a significant negative relationship between Age in years and the Response exclusively for the 250 Hz sound frequency (Table 3), indicating that younger children tended to draw larger circles compared to older participants whereas for other sound frequencies, drawn size did not change across age (Fig. 4). This result confirmed our initial hypothesis based on previous findings indicating that size drawing ability develops with age (Lange-Küttner, 1997, 2004, 2009).

ANOVAs on the linear mixed model for the variability in the response (i.e. raw data in cm^2 and z-transformed data) show a significant effect of the Pure tone frequency for all data sets (raw data in cm^2: F (4, 266.13) = 9.12; $p < 0.001$; z-transformed data: F (4, 268.00) = 6.06; $p < 0.001$). Regarding Age in years we observe a significant effect for the raw data in cm^2 (F (1, 66.88) = 6.79; $p = 0.011$) but no significant effect for the z-transformed data (F (1, 67.00) = 1.36; $p = 0.247$). For both data sets no significant interactions are reported (raw data in cm^2: F (4, 266.15) = 2.15; $p = 0.074$; z-transformed data: F (4, 268.00) = 0.61; $p = 0.649$). Post hoc analysis on the raw data in cm^2 and the z-transformed data show significant differences for all comparisons between Pure tone frequencies except for the non-significant comparison between 2000 and 5000 Hz (Table S2). Data are presented in Fig. 3 (right).

The correlation analysis (Spearman) for each sound frequency between size estimate and variability shows a significant positive correlation between participants' average size estimate across repetitions and variability (see Table 4) indicating that when drawing bigger circles participants tend to show higher variability. This pattern is in line with the Weber-Fechner law (Fechner, 1860; Weber, 1851), indicating that variability increases with drawing size magnitude for each sound frequency.

Participants were also asked to evaluate the size of the drawn shape on a scale from 1 to 10 (see Methods). Correlation analysis between the size estimate and the numerical rate assigned to each shape shows a positive correlation for each Pure tone frequency (Supplementary Table S3) confirming that children were able to translate the drawn size into symbolic representations.

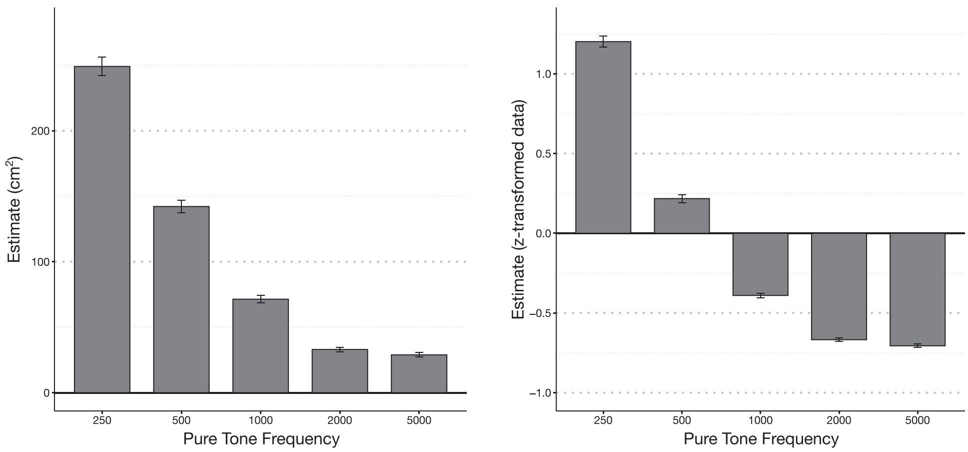

Fig. 2. Study 1 – Size estimate. Each bar represents average size estimate in cm^2 (left panel) and in z-transformed data (right panel). Error bars indicate standard error.

Table 1
Clinical details for visually impaired participants. Residual vision is expressed in LogMAR. NPL indicates no perception of light.

Participant ID	Age (y. o.)	Gender	Visual status	Pathology	Onset	Remaining Vision/Visual Acuity (LogMAR)
B01	11	M	Blind	Optic Nerve Glioma	LATE	Light perception
B02	6	F	Blind	Degenerative Maculopathy	LATE	Light perception
B03	13	F	Blind	Retinopathy of prematurity (ROP)	BIRTH	NPL
B04	13	F	Blind	Retinopathy of prematurity (ROP)	BIRTH	NPL
B05	14	F	Blind	Retinopathy of prematurity (ROP)	BIRTH	NPL
B06	7	F	Blind	Retinopathy	BIRTH	1.69
B07	7	M	Blind	Optic Atrophy, Osteopetrosis	BIRTH	1.69
B08	7	F	Blind	Microphthalmus (Left Eye), Anophthalmus (Right Eye)	BIRTH	1.69
LV01	6	M	Low vision	Bilateral Glaucoma	BIRTH	1.3
LV02	12	F	Low vision	Microphthalmus, Nistagmus	EARLY	1
LV03	7	M	Low vision	Leber Amaurosis	BIRTH	1
LV04	8	M	Low vision	Coloboma	BIRTH	1.3
LV05	12	F	Low vision	Retinitis Pigmentosa	EARLY	1.3
LV06	14	F	Low vision	Achromatopsia	EARLY	1
LV07	9	M	Low vision	Coloboma and microphthalmus	BIRTH	1
LV08	14	F	Low vision	Degenerative Maculopathy	EARLY	1
LV09	14	F	Low vision	Pilocytic astrocytoma	BIRTH	1
LV10	11	M	Low vision	Bilateral optic atrophy	BIRTH	0.69
LV11	10	M	Low vision	Albinism	BIRTH	1
LV12	9	M	Low vision	Congenital Cataract	LATE	1
LV13	6	M	Low vision	Albinism	BIRTH	1

Table 2
Age comparison.

Sighted			Blind			Low Vision		
ID	Age (y.o.)	Gender	ID	Age (y.o.)	Gender	ID	Age (y.o.)	Gender
S01	7	F	B02	6	F	LV13	6	M
S02	7	M	B06	7	F	LV01	6	M
S03	8	M	B07	7	M	LV03	7	M
S04	8	M	B08	7	F	LV04	8	M
S05	10	F				LV12	9	M
S06	10	M				LV07	9	M
S07	10	M				LV11	10	M
S08	12	F	B01	11	M	LV10	11	M
S09	13	M				LV05	12	F
S10	13	F	B03	13	F	LV02	12	F
S11	13	M	B04	13	F	LV06	14	F
S12	13	M	B05	14	F	LV08	14	F
S13	15	M				LV09	14	F

Table 3
Post-hoc (via emtrends) for the size estimate in cm^2 on the interaction between Age in years and Pure tone frequency in Study 1. * indicates $p < 0.01$.

Pure tone frequency	Age in years - trend	SE	df	lower.CL	upper.CL	t.ratio	*p*	
250	-17.44	5.36	88	-28.09	-6.80	-3.25	0.001	*
500	-0.69	5.37	89	-11.36	9.97	-0.12	0.897	
1000	1.80	5.37	89	-8.86	12.46	0.33	0.737	
2000	2.15	5.37	89	-8.51	12.82	0.40	0.689	
5000	0.63	5.35	88	-10.00	11.27	0.11	0.906	

3. Study 2

3.1. Methods

3.1.1. Participants

In Study 2, we investigated how visual impairment impacts the ability to associate drawn size with auditory pitch across development. As shown in Table 1, visual impairment is assessed according to the criteria established by the World Health Organization (World Health Organization, 2021): low vision is defined as a best-corrected visual acuity worse than 0.5 LogMAR but equal to or better than 1.3 LogMAR in the better eye. Blindness is defined as a best-corrected visual acuity worse than 1.3 LogMAR. 8 blind

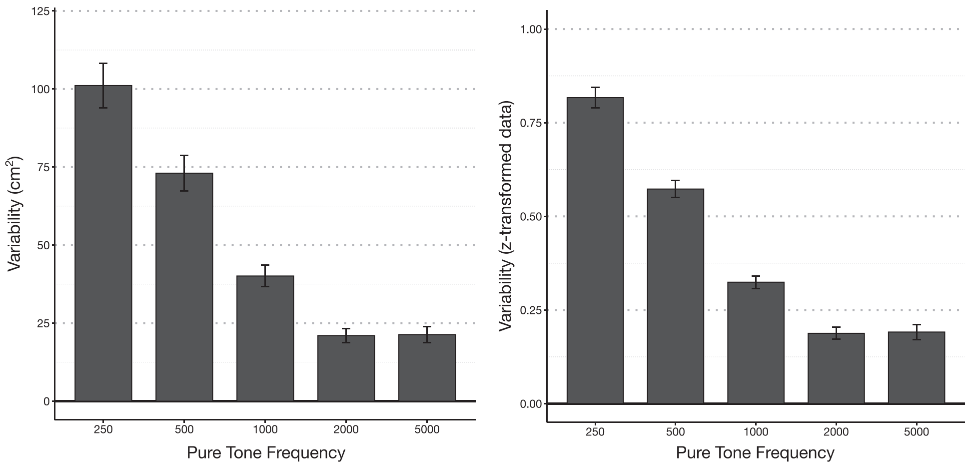

Fig. 3. Study 1 – Variability. Each bar represents the variability in the response averaged across subjects in cm^2 (left panel) and in z-transformed data (right panel). Error bars indicate standard error.

children (B01–08 in Tables 1 and 2; age range: 6–14 years, 79–171 months; 2 males; all right-handed), 13 low vision children (LV01–13 in Tables 1 and 2; age range: 6–14 years, 74–174 months; 8 males; all right-handed), and 13 age-matched sighted children with reported normal visual acuity (S01–13 in Table 2; age range: 7–15 years, 90–186 months; 9 males; all right-handed, except 1) participated. Visual acuity in visually-impaired children was assessed by ophthalmologists at rehabilitation center, David Chiossone Institute (Genoa, Italy). Age is not normally distributed for each group; thus, we ran non-parametric comparisons (i.e. Wilcoxon rank-sum test) that show no significant differences in terms of age between groups (low vision VS sighted: $W = 62.5$, $p = 0.41$; blind VS sighted: $W = 38$, $p = 0.32$; low vision VS blind: $W = 51$, $p = 0.84$). All participants had normal hearing. Visually impaired children were recruited from the rehabilitation center, David Chiossone Institute (Genoa, Italy). As reported by their rehabilitators, recruited children with low vision were able to draw simple and complex figures and all blind children followed an orientation and mobility training program that trained their general geometrical abilities such as simple shape recognition. Sighted children were recruited from the internal database of the Italian Institute of Technology (Genoa, Italy) to match the age of visually-impaired children and serve as a control group.

This study was carried out in accordance with the recommendations of the local health service, Comitato Etico, ASL 3 (Genoa, Italy) with written informed consent from all subjects or their legal representatives in accordance with the Declaration of Helsinki. The protocol was approved by the local ethical committee, Comitato Etico, ASL 3, (Genoa, Italy).

3.1.2. Stimuli

Stimuli were the same as described in Study 1 (see 2.1.2). In Study 2, we reduced the number of trials to avoid concentration loss in visually-impaired children. Thus, repetitions per frequency were reduced to 5, leading to 25 trials in total. The experiment consisted of one block of trials (duration: 10–15 min per block).

3.1.3. Task

The task was the same as described in Study 1 (see 2.1.2). Since blind children cannot take advantage of the visual feedback and low vision children could differentially benefit of visual feedback depending on the amount of residual vision, in Study 2 we blindfolded all children (including sighted participants) to allow comparisons between groups, thus the visual stimulus was not visible. As in Study 1, participants were asked to provide a rating of the drawn shape. However, when asked to perform the numeric task, some visually-impaired children (the youngest ones) showed distraction and refused to respond, thus we excluded these responses from the analysis. On each trial, the experimenter ensured participants provided a response within 30 s after sound presentation. In cases of no answer or participant's request, the experimenter repeated the trial and collected a response.

3.2. Data analysis

In Study 2, we tested the influence of the between-factor "Group" (with three levels: sighted, low vision, blind) and the within-factor "Pure tone frequency" (with five levels: 250, 500, 1000, 2000, 5000 Hz) on the size estimate. The variability of participant factor and the order of presentation were taken into account by modeling them as fixed random effects. The model is thus described by the following formula:

$$\text{RESPONSE} \sim \text{Group} \times \text{Pure tone frequency} + (1|\ \text{Participant}) + (1|\ \text{Trial Order})$$

where RESPONSE indicates the size estimate for each trial.

We conducted a similar analysis for the variability using a linear-mixed effects model described as follows:

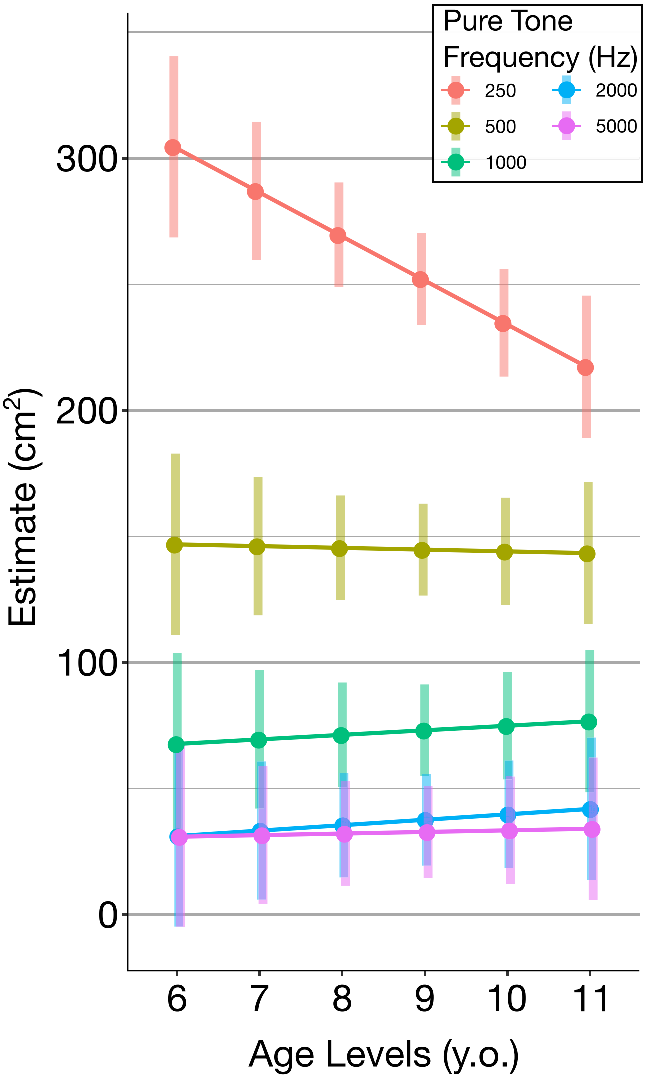

Fig. 4. Study 1 – Interaction age and pure tone frequency. The relationship between size in cm^2 and age is represented for each Pure tone frequency. Error bars represent standard error. Single age levels are represented for figurative reasons.

Table 4
Correlation analysis (Spearman) for each sound frequency between-participant average size estimate and variability across repetitions in Study 1.

Pure tone frequency (Hz)	Rho	*p*
250	0.72	< 0.0001
500	0.77	< 0.0001
1000	0.83	< 0.0001
2000	0.84	< 0.0001
5000	0.92	< 0.0001

All *p values are Bonferroni corrected.*

VARIABILITY ~ Group × Pure tone frequency + (1| Participant)

where VARIABILITY indicates the standard deviation across repetitions for each sound frequency and participant.

For both size estimate and variability, we performed ANOVA on the mixed model using Kenward-Roger's degrees of freedom approximation (Luke, 2017) to evaluate the significance of the model effects. Post hoc analysis was conducted by comparing contrasts using the emmeans package (Lenth et al., 2021). With this analysis we compute and contrast predicted probability distributions relative to each factor. We consider pairwise comparisons with $p < 0.05$ to be significant (Tukey's corrected *p*s are reported).

Since intermediate levels of compromised visual acuity would reveal the dynamics behind the development of the investigated crossmodal correspondences, we focused the analysis of Study 2 on the influence of residual visual acuity in LogMAR on performance. We thus tested the influence of the between-factor residual visual acuity in LogMAR (with 6 levels: 0; 0.69; 1; 1.3; 1.69; 2 LogMAR) and the within-factor "Pure tone frequency" (with 5 levels: 250, 500, 1000, 2000, 5000 Hz) on the size estimate. To allow comparison with sighted and blind children, we assigned a value of 0 LogMAR to sighted participants (who all reported normal visual acuity and did not wear glasses, as defined by the inclusion criteria) and a value of 2 LogMAR to blind participants who have no report of visual acuity in LogMAR, including participants who report no perception of light (NPL) and light perception only (B01–05 in Table 1). The variability of the participant factor is taken into account by modeling them as fixed random effects. The model is thus described by the following formula:

RESPONSE ~ LogMAR × Pure tone frequency + (1| Participant)

Since this analysis does not take into account the whole pattern of response, we ran additional analyses on the effect of residual visual acuity. We defined two possible response patterns: the predicted pattern, that is, large drawings in response to low pitch sounds and smaller drawings in response to high pitch sounds and a reversed/flat pattern, indicating an opposite or smaller association of sound frequency with drawn size. Specifically, for each participant, we performed a linear regression between the sound frequencies and the raw size estimate in cm^2 and extrapolated the slope: a negative slope indicates the predicted pattern, a positive slope indicates a reversed pattern. We then performed a one-way ANOVA on the slope, with residual vision in LogMAR as between-factor.

3.3. Results

Outlier exclusion led us to remove 5.21% of data points for the raw data in cm^2 and 4.96% of data points for the z-transformed data.

ANOVA results using the linear mixed model for each measurement of size estimation report a significant effect of the Pure tone frequency (raw data in cm^2: $F(4, 21.65) = 18.02$, $p < 0.001$; z-transformed data: $F(4, 21.92) = 47.47$, $p < 0.001$) and no significant effect of Group (raw data in cm^2: $F(2, 29.97) = 2.02$, $p = 0.149$; z-transformed data: $F(2, 30.00) = 0.42$, $p = 0.660$). We observe a trend for interaction between the two factors in the analysis of the raw data in cm^2 ($F(8, 718.99) = 1.90$, $p = 0.056$), which is confirmed by a significant interaction when considering z-transformed data ($F(8, 723.78) = 6.83$, $p < 0.001$). Results for Study 2 are shown in Fig. 5. Sighted participants show a pattern of results similar to those observed in Study 1; that is, they draw small circles in response to high pure tone frequencies and big circles in response to low pure tone frequencies. Low-vision participants do not show neat correspondences between the size of the drawn area as their size estimates are spread over the whole range of sizes. Interestingly, blind children follow a similar pattern to sighted participants. As reported in Table 5, post hoc analysis on both raw data in cm^2 and z-transformed data reveal significant differences between pure tone frequencies mostly in sighted participants (5 over 10 comparisons for the raw data in cm^2 and 7 over 10 comparisons for the z-transformed data), whereas significant differences between groups are present only for the pure tone frequency of 250 Hz when comparing z-transformed data from sighted with blind and low-vision children.

Since we observe an influence of the interaction between visual impairment and pure tone frequency on size estimate, we ran additional analyses to unveil the role of residual vision in the performance in terms of LogMAR. We ran ANOVA for each measurement of size estimation (i.e. raw data in cm^2 and z-transformed data) with residual visual acuity in LogMAR as between-factor (6 levels: 0; 0.69; 1; 1.3; 1.69; 2 LogMAR) and Pure tone frequency as within-factor (5 levels, i.e.: 250, 500, 1000, 2000, 5000 Hz). For all measurements of size estimation, this analysis confirms a significant effect of the pure tone frequency (raw data in cm^2: $F(4, 65.59) =$

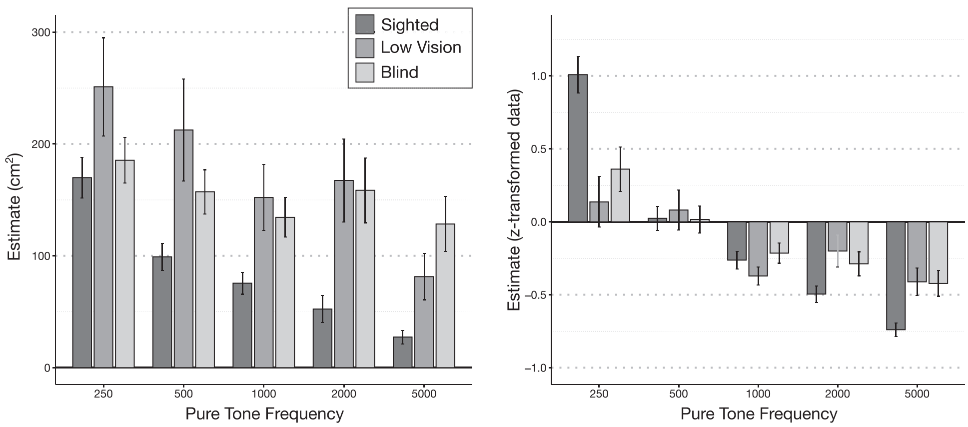

Fig. 5. Study 2 – Size estimate. The average size estimate in response to each sound frequency is represented for each group of participants. The left panel shows the size estimate in cm^2; the right panel represents the size estimate in z-transformed data. Dark gray bars represent data for age-matched sighted participants, medium gray bars show data for low-vision participants and light gray bars show data for blind participants. Error bars represent standard error.

Table 5
Post-hoc planned contrasts for the size estimate in cm^2 data between pure tone frequencies for each group in Study 2. * indicates p < 0.05; ** indicates p < 0.01; *** indicates p < 0.001; **** indicates p < 0.0001.

	Comparison	Estimate	SE	df	t.ratio	*p*	
Sighted	250 VS 500	69.73	23.50	127	2.96	0.029	*
	250 VS 1000	96.60	23.40	123	4.13	< 0.001	***
	250 VS 2000	115.69	23.60	127	4.90	< 0.0001	****
	250 VS 5000	143.12	23.40	125	6.10	< 0.0001	****
	500 VS 1000	26.87	23.10	119	1.16	0.773	
	500 VS 2000	45.96	23.30	122	1.97	0.286	
	500 VS 5000	73.38	23.20	121	3.15		*
	1000 VS 2000	19.09	23.20	119	0.82	0.922	
	1000 VS 5000	46.52	23.00	117	2.01	0.263	
	2000 VS 5000	27.43	23.30	121	1.17	0.763	
Blind	250 VS 500	33.08	29.40	260	1.12	0.792	
	250 VS 1000	99.06	29.20	255	3.39	0.007	**
	250 VS 2000	81.86	29.20	255	2.80	0.042	*
	250 VS 5000	160.18	29.40	260	5.44	< 0.0001	****
	500 VS 1000	65.98	29.40	260	2.24	0.165	
	500 VS 2000	48.78	29.40	260	1.65	0.461	
	500 VS 5000	127.10	29.60	266	4.29	< 0.001	***
	1000 VS 2000	-17.19	29.20	255	-0.58	0.976	
	1000 VS 5000	61.12	29.40	260	2.07	0.232	
	2000 VS 5000	78.32	29.40	260	2667.00	0.061	
Low Vision	250 VS 500	24.68	24.30	140	1.01	0.846	
	250 VS 1000	47.06	24.30	140	1.94	0.301	
	250 VS 2000	25.95	24.20	138	1.07	0.819	
	250 VS 5000	54.68	24.20	139	2.26	0.164	
	500 VS 1000	22.38	24.40	142	0.91	0.889	
	500 VS 2000	1.26	24.30	141	0.05	1.000	
	500 VS 5000	29.99	24.30	141	1.23	0.732	
	1000 VS 2000	-21.12	24.20	139	-0.87	0.907	
	1000 VS 5000	7.62	24.20	140	0.31	0.997	
	2000 VS 5000	28.73	24.10	138	1.19	0.756	

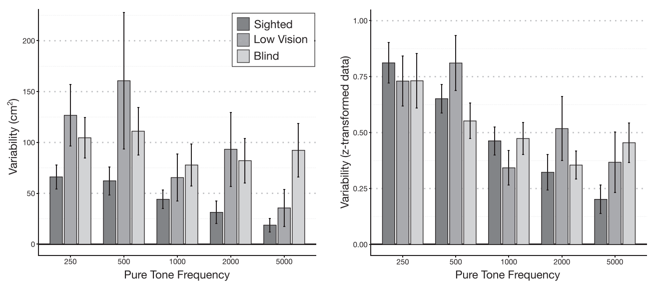

Fig. 6. Study 2 – Variability. The variability in response to each sound frequency is represented averaged across all participants. The left panel shows variability in cm^2, the right panel represents variability in z-transformed data. Error bars represent standard error.

6.77, $p < 0.001$; z-transformed data: F (4, 65.82) = 21.14, $p < 0.001$) but no significant effect given by residual visual acuity in LogMAR (raw data in cm^2: F (5, 26.97) = 17.84, $p = 0.149$; z-transformed data: F (5, 26.6) = 0.3, $p = 0.905$). We observe a significant interaction between Pure tone frequency and residual visual acuity (raw data in cm^2: F (20, 706.96) = 48.39, $p < 0.001$; z-transformed data: F (20, 710.49) = 8.75, $p < 0.001$). Considering that this analysis shows an interaction between residual vision and sound frequency, we hypothesize that the amount of residual vision may influence the pattern of drawn size in response to sound frequency. To test this hypothesis, we focus here on the pattern of responses given by participants to all sound frequencies; that is, the slope extrapolated from the linear regression between the size estimate and the sound frequencies for each participant (see Data Analysis for details and Fig. 7). We ran a one-way ANOVA on this parameter with residual vision in LogMAR as main factor (6 levels: 0; 0.69; 1; 1.3; 1.69; 2 LogMAR). The analysis shows a significant effect of residual vision for each measurement of size estimation, that is raw data in cm^2 (F (5, 27) = 2.95, $p = 0.02$, $\eta^2 = 0.35$) and z-transformed data (F (5, 27) = 3.34, $p = 0.01$, $\eta^2 = 0.38$).

For each measurement of response variability ANOVAs on the linear mixed model show a significant effect of Pure tone frequency (raw data in cm^2: $F(4, 119.06) = 6.89, p < 0.001$; z-transformed data: $F(4, 120) = 16.02, p < 0.001$) but no significant effect of Group (raw data in cm^2: F (2, 29.97) = 3.06, $p = 0.061$; z-transformed data: F (2, 30) = 0.29, $p = 0.752$) nor do we observe an interaction between the two factors (raw data in cm^2: $F(8, 119.06) = 1.49, p = 0.166$; z-transformed data: $F(8, 120) = 1.84, p = 0.074$) (Fig. 6).

4. Discussion

This study investigates 1) how crossmodal correspondences between auditory pitch and visual drawn size change with age in typically developing children and 2) how visual impairment affects the development of such association across childhood. In Study 1, we tested children from age 6–11 by asking them to draw a circle "as big as they heard the sound" and rate the drawn circle size using a number between 1 and 10. In Study 2, we used a similar protocol with visually-impaired children and adolescents; results were compared with a group of age-matched sighted participants. Results for both experiments show that sighted children already at the age of 6 associate high pure tone frequencies with smaller drawings and low pure tone frequencies with larger drawings, confirming our hypothesis on the early development of such crossmodal correspondences. We observed age-dependent modulation of the Audio-visual association exclusively for drawings in response to the sound frequency of 250 Hz, with older children drawing smaller shapes compared to younger participants. Such developmental pattern confirms our initial hypothesis indicating that drawing ability develops with age as previously observed in the literature (Lange-Küttner, 1997, 2004, 2009). The results of Study 2 suggest that the ability to associate the size of a drawn object with pure tones depends on the amount of residual vision thus supporting our hypothesis on the fundamental role of vision in the acquisition of such size associations across development. We discuss our results in terms of the role of multisensory (re-) calibration influencing crossmodal correspondences between pure tone frequencies and size.

Audio-visual crossmodal correspondences have been extensively studied; ample research demonstrates they are present in infants (Fernández-Prieto et al., 2015; Peña et al., 2011), children (Cuturi et al., 2019), adults and in animal models (Fitch & Hauser, 2003). In Study 1, we observed that sighted children aged 6–11 draw big circles in response to low pure tone frequencies and small circles in

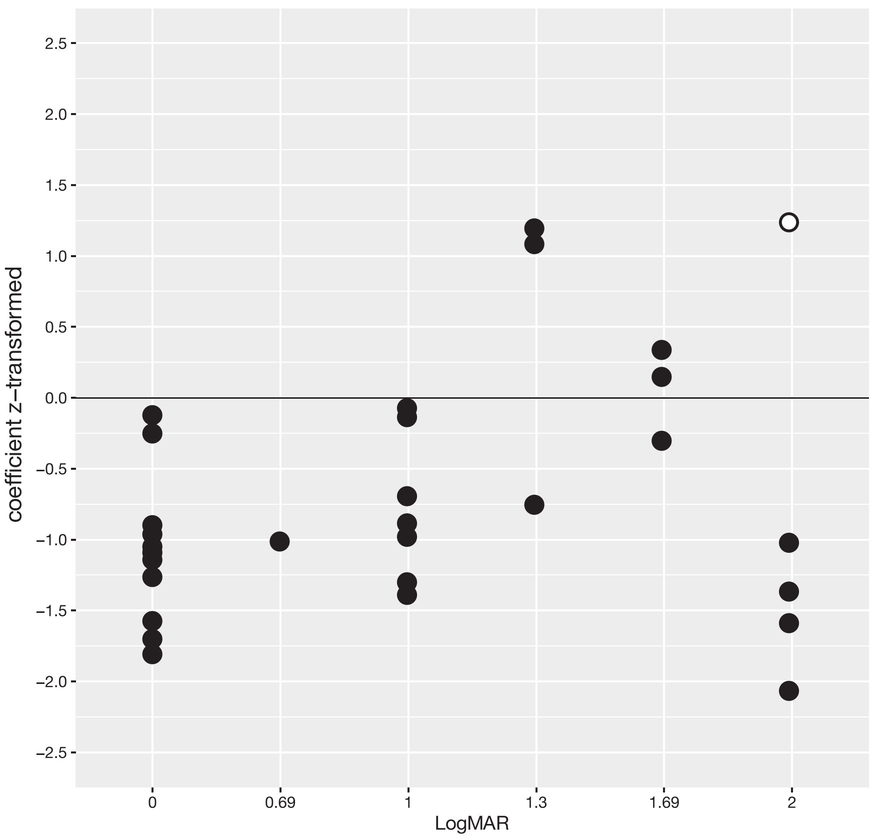

Fig. 7. Study 2 - Pattern of responses. The slope data for each participant (see 3.2 Data analysis of Study 2 for details) are plotted against the visual acuity LogMAR. The horizontal line represents the coefficient equal to 0, which means a flat pattern. The white data point represents the individual participant with a particular clinical history (see B01 in Table 1 and Discussion).

response to high pure tone frequencies. In a recent study (Cuturi et al., 2019) using a similar procedure, the authors tested crossmodal correspondences across childhood by asking children to identify the shape corresponding to a presented sound instead. The results report a developmental trend specific for the use of intermediate sizes in response to intermediate sound frequencies, in which older children are better able to take advantage of such precise and fine-grained stimuli compared to younger peers. These results are in contrast to what we observed here. However, in the present study, children were allowed to draw a circle with their dominant index finger corresponding to the heard sound while seeing the drawing. The absence of performance constraints in the response might explain the unobserved developmental trend in the present study. When asked to specifically choose the visual object associated with the perceived sound frequency, older children were capable of recognizing the scale and made better use of the intermediate visual object. In the case of free drawing, this advantage seems attenuated, as younger and older children behave similarly. It can be argued that the nature of the response, being continuous and thus more prone to biases and variable error, is less effective in providing a measurement of fine tunings of crossmodal correspondences compared to an identification task where responses are discrete. A possible explanation behind the differences between the two tasks may rely on categorical processes involved in the identification task compared to the reproduction task. On one hand, the presence of visual objects in the response phase of the identification task might involve bottom-up processes in the drawing phase that directly allow participants to use intermediate size levels to define the perceived sound frequency. In particular, the emergence of a categorical membership for auditory stimuli may be related to a learning process that shapes the neural response to sound properties across development (Bizley & Cohen, 2013) likely defining finer perceptual encoding of auditory features as age increases. Such processes might then be activated by the presence of a categorical classification in the response phase. On the other hand, drawing a circle in response to the sound might be based on top-down processes that show the use of crossmodal correspondences, which might not include the ability to provide fine-grained judgments. Overall, the absence of constraints in the response phase of the drawing task — that is, the absence of a forced-choice requirement as in an identification task — reinforces and confirms the hypothesis that crossmodal correspondences are spontaneously present at the age of 6. Along these lines, the employment of goal-directed motor tasks within an egocentric frame of reference has been shown to be strongly related to the definition of object-related such as size rather than spatial properties (Creem & Proffitt, 2001; Milner & Goodale, 1995).

The impact of visual impairment on the association of size with sound frequency seems to rely on the progressive decay of visual acuity across the life span. Recent findings focusing on the role of visual experience in sound-size correspondences have shown no differences between blind and sighted adults when haptic information is used to encode object size (Hamilton-Fletcher et al., 2018). Vision might not be indispensable to develop such associations; however, other senses, such as haptic might be involved in the generation of crossmodal correspondences relative to size (Hamilton-Fletcher et al., 2018). Supporting this view, developmental studies have shown that children tend to rely more on haptic sense rather than vision when they need to judge object size compared to orientation judgment, where instead vision is the leading sensory modality (Gori et al., 2008). Although it is known from the literature that haptic size thresholds do not depend on residual visual acuity (Gori et al., 2010), we observed an influence of residual vision on drawn size when the latter is expressed in response to audition, a sensory modality not intrinsically related to size. Specifically, we observe that sighted participants and participants with mild low vision (LogMAR between 0.69 and 1) show a strong association of low pure tones with big drawings and high pure tones with small drawings. In contrast, it is only from more severe visual impairment (LogMAR > 1.3) that we start to observe some participants with an opposite pattern of responses, with low pure tones associated with small drawings and high pure tones with big drawings. Children with a total absence of residual visual acuity (LogMAR $= 2$) show a pattern similar to sighted and mild low vision children; namely, they tend to associate low sound frequencies with large drawings and high sound frequencies with small drawings. Although the low sample size and the high variability among participants prevent us from drawing definitive conclusions, we speculate on the mechanisms that might underlie this pattern of results. One possibility regards the presence of a dynamic re-calibration process, taking place as visual impairment advances. Such processes have been reported showing sensory remapping in the Audio-visual processing of space (Garcia et al., 2017). In sighted children, pure tone-size associations are consolidated thanks to the crossmodal matching of visually (and haptically) defined size and the corresponding perceived sound. Similarly, low-vision children with higher visual acuity might still benefit from these processes as their reliability on multisensory cues is less compromised. The impoverishment of visual experience in some of the severe low vision children seems to overall affect their ability to translate audition in the reproduction of size. In this sense, our results may support the idea that visual acuity has a role in the consolidation of pure tone-size correspondences in visually-impaired children. In case of severe visual impairment, the overall absence of an auditory-size pattern suggests that severe reduction of visual acuity prevents the development of such association; however, this is not true for all blind participants. We indeed observed that blind participants with no residual visual acuity (B02–05, see Table 1) show a pattern of responses similar to sighted participants. Only one blind participant showed a strong reversed pattern of responses (B01, Table 1 and white data point in Fig. 7) and differed from others for his particular clinical history. Compared to other blind children, visual loss of this participant increased across the lifespan reaching complete blindness not long before the moment of testing ($<$ 1 year). This observation may suggest that a complete lack of vision for a prolonged period can induce compensatory mechanisms that allow totally blind individuals to process size through audition similarly to sighted peers. Such observation is in line with the aforementioned presence of audio-haptic crossmodal correspondences in blind and sighted adults (Hamilton-Fletcher et al., 2018). Along these lines, life experience through the remaining senses may enhance the ability to take advantage of sensory modalities; thus, auditory information becomes functional to process objects features usually transmitted through vision. On the other hand, severe low-vision children with lower visual acuity might stress their residual vision instead of increasing sensory reliability of other modalities that are not intrinsically linked to size, such as audition. With a reduction of visual acuity, it might be possible that the recalibration process between pure tone frequencies and size is temporary compromised due to an anchoring on residual vision that does not allow crossmodal correspondence to emerge. In this context, haptic processing might be critical, as touch gradually becomes the most reliable sense to process size and likely to allow the (re-) gaining of size reproduction through auditory information. It is therefore suggested to investigate the role of haptic information in such crossmodal associations specifically in relationship to visual impairment of different degrees. In this context, deeper investigation on the role of age onset or the duration of blindness across the lifespan would throw light on the sensory-loss compensatory processes underlying size perception. It would indeed be desirable to expand the results of the present research by increasing the sample size of visually impaired children to trace a more detailed developmental path for the emergence of crossmodal correspondences possibly by including longitudinal follow-ups.

Alternatively, it can be argued that children with low vision and blindness may exhibit different motor patterns regarding the use of fingers to draw the response. According to this view, children with typical development and low vision might have gained the ability to use their fingers and other tools (e.g. pens/pencils and touchscreens) to draw and interact very similar to the procedure adopted in this study, whereas children with severe low vision and blindness might have developed with different sets of psychomotor patterns and might not be familiar with the use of a touchscreen. While we cannot exclude a potential effect given by such intrinsic and idiosyncratic differences in the drawing experience and ability, were such a scenario true, the observation of a clear pattern of crossmodal correspondences in the blind group of participants would remain unexplained. Further investigation may then focus on the relationship between psychomotor abilities and the association of pure tone frequency with size, possibly by testing participants using tasks that do not require a motor response and/or by assessing psychomotor abilities prior to administration of the task.

However, the association between variability in drawn size and pure tone frequency seems to be present regardless of visual impairment. Specifically, all participants show comparable variability in response to sound frequencies; that is, a pattern of variability associated with the presented sound frequency. In other words, regardless of the drawn size, participants are more variable in giving their response when asked to draw circles after listening to low pure tone frequencies and less variable in response to high pure tone frequency. To better understand this result, a few considerations can be made. First, in Study 2, participants who showed the predicted pattern of association between sound frequency and drawn size are more than those who showed a reversed/flat pattern. Second, in Study 1, we observe a linear relationship between drawn size and variability associated with drawing, which is in line with the Weber-Fechner law; that is, perceived variability increases with magnitude of the stimulus, in this case, the drawing. Given these considerations, our results hint towards an overall association of sound frequency with variability in drawing, which is likely to be more primitive than the relationship between sound frequency with drawing size. The latter process might require a longer experience of visual loss to fully emerge, as in total blind participants.

5. Conclusions

Altogether, the results from Study 1 to 2 provide novel insights relative to sound frequency-size crossmodal correspondences. On one hand, sighted children show such association as early as the age of 6, indicating that these correspondences are present early in life, as previously observed. On the other, the level of residual visual capacity in visually-impaired children seems to be involved in the process of manifesting pure tone frequency-size correspondences. Size-pure tone frequency associations might depend on how sensory information is processed, in particular, on how participants rely on different sensory modalities. Investigation of crossmodal correspondences relative to pure-tone frequency and size is of interest for the development of pedagogical tools used to teach size comparisons, since auditory pure tone frequency can provide an additional cue to be used to discriminate sizes and thus may enhance multisensory teaching methods in sighted (e.g., Gori, Volpe, Cappagli, Volta, & Cuturi, 2021) and visually impaired children. In this study, we observed that the correspondence between size and sound in visually-impaired children depends on the level of visual impairment, suggesting that this association manifests when visual impairment reaches severe levels when the need to use remaining sensory modalities is indispensable. Further research may consider several aspects that emerged from the current study. While we report visual impairment onset by following the WHO classification (World Health Organization, 2021), more precise information relative to the duration of blindness in terms of years or months would throw light on the dynamics that underlie the emergence of size-sound frequency associations. Most children enrolled in this study had degenerative pathologies that induce progressive visual impairments making impossible to define precise visual impairment onset. Moreover, it is important to underline that children included in this study are enrolled in a rehabilitation program aimed at increasing their autonomy and their general sensorimotor and cognitive abilities. Especially in the case of totally blind children, orientation and mobility training leads to improvements into two main contexts. First, it provides them with a means to spatially navigate and orient themselves based on their remaining sensory modalities; they indeed learn specific techniques that allow them to understand traffic direction or openness of the surrounding space. At the same time, such training stresses their abilities to comprehend geometrical principles via sensorimotor stimulation; for instance, by drawing shapes while moving. Although these aspects may suggest that rehabilitation programs focusing on sensorimotor processes could have a role in the emergence of crossmodal association between sound frequency and drawn size, a control group that did not undergo such a training program would be needed. Such limitation is unlikely to be overcome, though, as most visually-impaired children follow rehabilitation programs to improve their cognitive and sensorimotor abilities. Nevertheless, these considerations, and the results presented in this study, point towards a fundamental role of the use of remaining senses to understand the surrounding space, which may lead to improved multisensory abilities such as the crossmodal association object of this study. Along these lines, the association of sound and reproduced size investigated here is of great interest in the context of therapeutic interventions for visually impaired children. Such association may be learned by or reinforced in visually-impaired children to increase their reliability on other senses than vision to encode concepts as an object's size.

CRediT authorship contribution statement

Luigi F. Cuturi: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Giulia Cappagli**: Conceptualization, Investigation, Writing – review & editing. **Alessia Tonelli**: Conceptualization, Methodology, Writing – review & editing, Formal analysis. **Elena Cocchi**: Recruitment of visually impaired children. **Monica Gori**: Conceptualization, Methodology, Writing – review & editing, Supervision.

Funding

This work was partially supported by the EU- H2020-ICT Project WeDRAW. WeDRAW has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 732391. The research is partially supported by the MYSpace project (PI Monica Gori), which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 948349). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We thank all participants who took part in the study, the Istituto Don Bosco (Genoa, Italy) and the David Chiossone Institute (Genoa, Italy) for their kind availability. We also thank Elisa Freddi for assistance with data collection. The content of this publication is the sole responsibility of the authors. The European Commission or its services cannot be held responsible for any use that may be made of the information it contains.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cogdev.2021.101125.

References

Amadeo, M. B., Campus, C., & Gori, M. (2019). Time attracts auditory space representation during development. *Behavioural Brain Research, 376*(June), Article 112185. https://doi.org/10.1016/j.bbr.2019.112185

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience, 14*(10), 693–707. https://doi.org/10.1038/nrn3565

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189–201. https://doi.org/10.1037/0012-1649.41.6.189

Cappagli, G., Cocchi, E., & Gori, M. (2015). Auditory and proprioceptive spatial impairments in blind children and adults. *Developmental Science*, 1–12. https://doi.org/10.1111/desc.12374

Cappagli, G., & Gori, M. (2016). Auditory spatial localization: Developmental delay in children with visual impairments. *Research in Developmental Disabilities, 53–54*, 391–398. https://doi.org/10.1016/j.ridd.2016.02.019

Creem, S. H., & Proffitt, D. R. (2001). Defining the cortical visual systems: "What", "Where", and "How". *Acta Psychologica, 107*(1–3), 43–68. https://doi.org/10.1016/S0001-6918(01)00021-X

Cuturi, L. F., Tonelli, A., Cappagli, G., & Gori, M. (2019). Coarse to fine Audio-visual size correspondences develop during primary school age. *Frontiers in Psychology, 10*(September), 1–10. https://doi.org/10.3389/fpsyg.2019.02068

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science, 25*(6), 1256–1261. https://doi.org/10.1177/0956797614528521

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition, 114*(3), 405–422. https://doi.org/10.1016/j.cognition.2009.10.013

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433. https://doi.org/10.1038/415429a

Evans, K. K., & Treisman, A. (2011). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*. https://doi.org/10.1167/10.1.6

Fechner, G. T. (1860), *Vol. 2. Elemente der psychophysik*. Breitkopf u. Härtel.

Fernández-Prieto, I., Navarra, J., & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development*. https://doi.org/10.1016/j.infbeh.2014.12.008

Fitch, W. T., & Hauser, M. D. (2003). Unpacking "honesty": Vertebrate vocal production and evolution of acoustic signals. *Acoustic Communication, 1990*, 65–137. https://doi.org/10.1007/0-387-22762-8_3

Fooks, N., Hadad, B.-S., & Rubinsten, O. (2021). Nonsymbolic-magnitude deficit in adults with developmental dyscalculia: evidence of impaired size discrimination but intact size constancy. *Psychological Science*, Article 0956797621995204.

Garcia, S. E., Jones, P. R., Reeve, E. I., Michaelides, M., Rubin, G. S., & Nardini, M. (2017). Multisensory cue combination after sensory loss: Audio-visual localization in patients with progressive retinal disease. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000344

Gori, M. (2015). Multisensory integration and calibration in children and adults with and without sensory and motor disabilities. *Multisensory Research, 28*(1–2), 71–99. https://doi.org/10.1163/22134808-00002478

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18*(9), 694–698. https://doi.org/10.1016/j.cub.2008.04.036

Gori, M., Giuliana, L., Sandini, G., & Burr, D. (2012). Visual size perception and haptic calibration during development. *Developmental Science, 15*(6), 854–862. https://doi.org/10.1111/j.1467-7687.2012.2012.01183.x

Gori, M., Sandini, G., Martinoli, C., & Burr, D. (2010). Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration. *Current Biology, 20*(3), 223–225. https://doi.org/10.1016/j.cub.2009.11.069

Gori, M., Volpe, G., Cappagli, G., Volta, E., Cuturi, L.F. (2021). Embodied multisensory training for learning in primary school children, in: 2021 IEEE International Conference on Development and Learning (ICDL). ⟨https://doi.org/10.1109/ICDL49984.2021.9515596⟩.

Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology, 22*(5), 383–388. https://doi.org/10.1016/j.cub.2012.01.004

Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition, 175*(February), 114–121. https://doi.org/10.1016/j.cognition.2018.02.023

Honoré, N., & Noë, M. P. (2016). Improving preschoolers' arithmetic through number magnitude training: The impact of non-symbolic and symbolic training. *PLoS ONE, 11*(11), 5–7. https://doi.org/10.1371/journal.pone.0166685

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological Science, 15*(1), 60–64. https://doi.org/10.1111/j.0963-7214.2004.01501010.x

Jaekl, P., Soto-Faraco, S., & Harris, L. R. (2012). Perceived size change induced by audiovisual temporal delays. *Experimental Brain Research, 216*(3), 457–462. https://doi.org/10.1007/s00221-011-2948-9

Lange-Küttner, C. (1997). Development of size modification of human figure drawings in spatial axes systems of varying complexity. *Journal of Experimental Child Psychology, 66*(2), 264–278. https://doi.org/10.1006/jecp.1997.2386

Lange-Küttner, C. (2004). More evidence on size modification in spatial axes systems of varying complexity. *Journal of Experimental Child Psychology, 88*(2), 171–192. https://doi.org/10.1016/j.jecp.2004.02.003

Lange-Küttner, C. (2009). Habitual size and projective size: The logic of spatial systems in children's drawings. *Developmental Psychology, 45*(4), 913–927. https://doi.org/10.1037/a0016133

Lenth, R., Buerkner, P., Herve, M., Love, J., Riebl, H., Singmann, H. (2021). Package "emmeans." ⟨https://doi.org/doi:10.1080/00031305.1980.10483031⟩.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Milner, D., & Goodale, M. (1995). The visual brain in action. In *The Visual Brain in Action* (Vol. 9780198524). Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198524724.001.0001

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/10.3758/CABN.4.2.133

Nag, S., & Snowling, M. J. (2013). Children's reading development: Learning about sounds, symbols, and cross-modal mappings. *Cognition and brain development: Converging evidence from various methodologies* (pp. 253–270). American Psychological Association.

Orchard-Mills, E., Van der Burg, E., & Alais, D. (2013). Amplitude-modulated auditory stimuli influence selection of visual spatial frequencies. *Journal of Vision, 13*(3). https://doi.org/10.1167/13.3.6 (6–6).

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology, 114*(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004

Parise, C. V., & Spence, C. (2009). "When birds of a feather flock together": synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One, 4*(5), 5664. https://doi.org/10.1371/journal.pone.0005664

Parise, C. V. (2015). Crossmodal correspondences: Standing issues and experimental guidelines. *Multisensory Research, 29*(October), 7–28. https://doi.org/10.1163/22134808-00002502

Peña, M., Mehler, J., & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychological Science, 22*(11), 1419–1421. https://doi.org/10.1177/0956797611421791

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: Children's integration of auditory and haptic information is suboptimal. *Developmental Science, 17*(3), 376–387. https://doi.org/10.1111/desc.12127

Scheller, M., Proulx, M. J., De Haan, M., Dahlmann-Noor, A., & Petrini, K. (2021). Late-but not early-onset blindness impairs the development of audio-haptic multisensory integration. *Developmental Science, 24*(1), 13001.

Silk, A. M. J., & Thomas, G. V. (1988). The development of size scaling in children's figure drawings. *British Journal of Developmental Psychology, 6*(3), 285–299. https://doi.org/10.1111/j.2044-835x.1988.tb01101.x

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Spence, C. (2019). On the relative nature of (pitch-based) crossmodal correspondences. *Multisensory Research, 32*(3), 235–265. https://doi.org/10.1163/22134808-20191407

Tivadar, R. I., Rouillard, T., Chappaz, C., Knebel, J. F., Turoman, N., Anaflous, F., & Murray, M. M. (2019). Mental rotation of digitally-rendered haptic objects. *Frontiers in Integrative Neuroscience, 13*(March), 1–9. https://doi.org/10.3389/fnint.2019.00007

Tonelli, A., Cuturi, L. F., & Gori, M. (2017). The influence of auditory information on visual size adaptation. *Frontiers in Neuroscience, 11*(OCT). https://doi.org/10.3389/fnins.2017.00594

Vuontela, V., Steenari, M., & Koivisto, J. (2003). Audiospatial and visuospatial working memory in 6–13 year old school children. *Learning & Memory, 10*, 74–81. https://doi.org/10.1101/lm.53503.which

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Weber, E.H. (1851). Der Tastsinn und das Gemeingefühl (1846). Published as Off-Print from R. Wagner's Handwörterbuch Der Physiologie.

Wiedenbauer, G., & Jansen-Osmann, P. (2008). Manual training of mental rotation in children. *Learning and Instruction, 18*(1), 30–41. https://doi.org/10.1016/j.learninstruc.2006.09.009

World Health Organization, (WHO). (2021). Blindness and vision impairment. ⟨https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment⟩.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Perception-action dissociations of a walkable Müller-Lyer configuration. *Psychological Science, 11*(3), 239–243. https://doi.org/10.1111/1467-9280.00248